DOCUMENT RESUME

ED 412 451 CG 028 080

AUTHOR Smith, Douglas K.; Buckley, Sharon; Shine, Agnes E.
TITLE WISC-III/WISC-R Relationships in Native Alaskan Students.
PUB DATE 1996-03-00
NOTE 16p.; Paper presented at the Annual Meeting of the National Association of School Psychologists (Atlanta, GA, March 12-16, 1996).
PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Adolescents; *Alaska Natives; Children; Comparative Analysis; *Construct Validity; Elementary Secondary Education; *Intelligence Tests; Longitudinal Studies; Measures (Individuals); Minority Group Children; *Test Validity
IDENTIFIERS *Wechsler Intelligence Scale for Children III; Wechsler Intelligence Scale for Children (Revised)

ABSTRACT

Since the introduction of the Wechsler Intelligence Scale for Children Third Edition (WISC-III), practitioners have raised questions concerning the equivalency of scores on the WISC-III and WISC Revised (WISC-R) in the re-evaluation of students receiving special education services. Since studies with minority students are lacking in this area, a study was designed to examine the relationship between WISC-III and WISC-R scores using a sample of Native Alaskan students being re-evaluated for possible continuation of special education services. Subjects included 47 Native Alaskan students, with a mean age of 9.64 years at initial evaluation with the WISC-R and a mean age of 12.39 years at re-evaluation with the WISC-III. Results of the study suggest stability of performance in the verbal areas of the Wechsler with nonsignificant differences between WISC-III and WISC-R mean subtest scores on Information, Similarities, Arithmetic, Vocabulary, and Digit Span. Verbal intelligence scores were 4.64 points higher on the WISC-III although previous studies found the scores to be consistent. Differences in scores were most likely to occur on the Performance Scale and least likely to occur on the Verbal Scale. (RJM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 412 451

WISC-III/WISC-R Relationships in Native Alaskan Students

Douglas K. Smith
Sharon Buckley
University of Wisconsin-River Falls

Agnes E. Shine
Barry University

BEST COPY AVAILABLE

Paper presented at the Annual Meeting of the National Association of School Psychologists, Atlanta, March 1996.

CGO28080

PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
D. Smith
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC).

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

ABSTRACT

Since the introduction of the WISC-III, practitioners have raised questions concerning the equivalency of scores on the WISC-III and WISC-R in the re-evaluation of students receiving special education services. There have been numerous reports of discrepancies in scores ranging from 10 to 20 points which is considerably larger than the expected five to seven point difference described in the WISC-III Manual. A recent review of WISC-III/WISC-R comparison studies found an average difference of 5.69 points (weighted for sample size) between the Full Scale IQ of the two tests with the WISC-III scores being lower. The majority of these 22 studies utilized nonminority students with learning disabilites. Since studies with minority students are lacking, this study was designed to examine the relationship between WISC-III and WISC-R scores using a sample of Native Alaskan students being re-evaluated for possible continuation of special education services. Subjects were 47 Native Alaskan students with a mean age of 9.64 years at initial evaluation with the WISC-R and a mean age of 12.39 years at re-evaluation with the WISC-III. Pearson product moment correlations ranged from .66 on the Performance Scale to .79 on the Verbal Scale. T-tests for related samples were significant ($p < .001$) for all global scales with WISC-III scores higher than WISC-R scores (mean difference of 5 to 12 points). Subtest analyses and the implications of the study for school psychology practice are discussed.

Since the introduction of the Wechsler Intelligence Scale for Children-III (WISC-III; Wechsler, 1991), practitioners have raised questions concerning the equivalency of scores on the two instruments in the re-evaluation of students receiving special education services. There have been numerous reports of discrepancies in scores ranging from 10 to 20 points, which is considerably larger than the expected five to seven point difference described in the WISC-III manual.

The expected difference in WISC-R/WISC-III scores was based on a study reported in the WISC-III Manual in which 206 children, ages 6 through 16 years, were administered the WISC-III and WISC-R in counterbalanced order wtih a median interval between testings of 21 days. Correlations between the WISC-III and WISC-R for the Verbal IQ (VIQ), Performance IQ (PIQ), and Full Scale (FSIQ) were .90, .81, and .89, respectively. Differences in mean scores were 2.4 points, 7.4 points, and 5.3 points, respectively, for the VIQ, PIQ, and FSIQ with WISC-III scores lower on each scale. An additional study, utilizing a clinical sample of 104 students, is also reported in the WISC-III Manual. This sample ranged in age for 7 to 14 years of age, was predominantly male and included "57% children with various learning and reading disabilities, 35% children with Attention-Deficit Hyperactivity Disorder, and 8% children with depression or anxiety disorders" (Wechsler, 1991, p. 199). Corrected correlations between the WISC-R and WISC-III VIQ, PIQ and FSIQ were .86, .73, and .86, respectively. WISC-III scores were lower (5.4 points for the VIQ, 5.1 points for the PIQ and 5.9 points for the FSIQ).

Two other studies involving WISC-R/WISC-III comparisons are reported in the WISC-III Manual. One study involved 23 children previously identified as

gifted. The mean interval between the previously administered WISC-R and the current WISC-III was 13 months. Mean scores on the VIQ, PIQ, and FSIQ were 5.8, 1.1, and 4.9 points lower on the WISC-III. A similar study with 28 students diagnosed as mildly mentally retarded is also reported. The mean test interval between the previously administered WISC-R and the current WISC-III was 2 years, 2 months. Mean WISC-III IQ scores were 8.9, 6.8, and 8.9 points lower than the mean WISC-R scores on the VIQ, PIQ, and FSIQ, respectively. While clinical samples were used in the WISC-III/WISC-R validity studies described in the manual, sample sizes were limited for students receiving special education services and, except for the students with severe conduct disorders, samples were primarily White.

Since publication of the WISC-III, several studies comparing WISC-R/WISC-III performance have been completed. Bolen, Aichinger, Hall, and Webster (1993) utilized a sample of 61 learning impaired students, due for re-evaluation of their special education placement. Each student was administered the WISC-III and comparisons were made with the previously administered WISC-R. Substantial differences were found in WISC-R and WISC-III VIQ (5.20 points), PIQ (9.21 points), and FSIQ (7.95 points) with mean WISC-III scores lower on all three scales. Potter and Wessel (1993) used a sample of 188 students being re-evaluated for continued special education services in the Milwaukee, Wisconsin public schools. The sample included all re-evaluations completed between November 1, 1992 and December 12, 1992 and the current WISC-III scores were compared to the previously obtained WISC-R scores. Differences were found on the VIQ (6.07 points), PIQ (8.33 points), and the FSIQ (7.72 points) with mean WISC-III scores lower than the WISC-R scores on all three scales. Smith, Stovall, and Geraghty (1995), used a

sample of 300 Wisconsin students being re-evaluated for possible continuation of special education services during the 1992-93 academic year. Current WISC-III scores were compared to the previously obtained WISC-R scores. Differences were found on the VIQ (3.65 points), PIQ (5.69 points), and FSIQ (5.03 points) with mean WISC-III scores lower than the WISC-R scores on all three scales. Meanwhile, Slate and Saarnio (1995) used a sample of 257 students undergoing psychoeducational re-evaluations in Mississippi and found mean differences of 5.8 points, 12.26 points, and 7.2 points on the VIQ, PIQ, and FSIQ, respectively. Finally, Weiss (1995) reviewed 22 published and unpublished WISC-III/WISC-R comparison studies and found an average difference of 5.69 points (weighted for sample size) between the FSIQ of the two tests with the WISC-III scores being lower. The majority of these studies utilized nonminority students with learning disabilities.

Purpose of the Study

The studies completed to date have produced a mixed pattern of results in comparing WISC-R/WISC-III performance. Sample sizes have been limited in some studies and narrowly defined clinical samples have been used in others. In addition, the studies have focused on nonminority students. Consequently, the generalizability of the results to other populations is limited. Therefore, the purpose of this study was to examine the relationship between the WISC-III and WISC-R scores using a sample of Native Alaskan students being re-evaluated for possible continuation of special education services.

Method

Subjects were 47 Native Alaskan students being re-evaluated for possible continuation of special education services during the 1993-94 academic year. Data collected included WISC-III/WISC-R subtest and global scores, sex of the

student, race/ethnicity of the student, community size, student age, initial disability of the student, type of program in which the student was enrolled (resource, self-contained, etc.), and the placement decision following the re-evaluation with the WISC-III. Data were coded and double-checked by school psychology graduate students prior to data analysis.

Students ranged in age from 6 years, 0 months to 13 years, 9 months (mean age of 9 years, 8 months) at initial testing with the WISC-R and from 7 years, 2 months to 16 years, 5 months (mean age of 12 years, 5 months) at the time of re-evaluation with the WISC-III. The sample was composed of 31 males and 16 females. Services for students with learning disabilities were being provided for 62% of the sample after the initial evaluation and for 79% of the sample after the re-evaluation. Speech/language services were being provided for 21% of the sample after the initial evaluation and for none of the sample after the re-evaluation. Services for students with mental retardation remained stable at 17% and 13%, respectively. Three students (6%) exited special education after the re-evluation. The entire sample was from communities of less than 2,000 residents.

Results and Discussion

Mean standard scores on the WISC-R global scales ranged from 71.02 (VIQ) to 80.68 (PIQ). For the WISC-III, the range was from 75.66 (VIQ) to 92.47 (PIQ). The mean difference in scores was 4.64 points on the VIQ, 11.79 points on the PIQ, and 8.00 points on the FSIQ, with higher scores on the WISC-III than the WISC-R. These differences in mean global scale socres were analyzed by t-tests for related samples and produced statistically significant results for the VIQ ($t = 4.90$, $p < .001$); for the PIQ ($t = 6.44$, $p < .001$); and for the FSIQ ($t = 6.78$, $p < .001$). The magnitude of the mean difference

in scores is similar to the difference reported in previous studies. However, the direction of the difference on all three global scales is different with WISC-III scores higher than WISC-R scores. These results are presented in Table 1.

Insert Table 1 about here

Mean WISC-R subtest scores ranged from 5.14 (Comprehension) to 7.89 (Picture Completion), while mean WISC-III subtest scores ranged from 5.10 (Arithmetic) to 9.21 (Object Assembly). The mean difference in subtest scores ranged from -.47 (Arithmetic) to +2.10 (Block Design). WISC-III subtest means were higher than WISC-R subtest means for all subtests except Arithmetic and Vocabulary. The differences in mean subtest scores were analyzed by t-tests for related samples and produced statistically significant results ($p < .05$) for these subtests: Comprehension, Picture Completion, Picture Arrangement, Block Design, Object Assembly, and Coding. These results are presented in Table 2.

Insert Table 2 about here

Pearson product moment correlations were calculated for the global scales of the two tests and produced these results: for VIQ, $r = .79$ ($p < .001$); for PIQ, $r = .67$ ($p < .001$); and for FSIQ, $r = .76$ ($p < .001$). These results suggest that the VIQ and the FSIQ scales of the two instruments are highly related and that the difference in mean scores, although statistically significant, may have limited practical significance with the exception of

the difference in PIQ scores which is much larger (approaching one standard deviation). The most important finding, however, is that with all three global scales the WISC-III scores were higher than the WISC-R scores for this sample of Native Alaskan students. This result is in contrast to other studies with mostly White samples in which WISC-III scores were lower than WISC-R scores. Pearson product moment correlations were calculated on the mean subtest scores of the WISC-III and WISC-R. Significant correlations ($p < .05$) ranging from .31 (Block Design) to .57 (Arithmetic) were obtained for all subtests except Information, Similarities, and Picture Arrangement. The mean verbal subtest correlation was .37. while the mean performance subtest correlation was .34. Thus, the subtests of the two instruments are moderately related. These results are presented in Table 3.

Insert Table 3 about here

The magnitude of the mean differences in scores on the VIQ and FSIQ are consistent with previous studies of students with learning disabilities. However, the direction of the difference is not consistent. In the current study mean WISC-III scores were higher than the previously obtained WISC-R scores. Only one previous study (Graf & Hinton, 1994) had reported a higher WISC-III mean than WISC-R mean on the VIQ and the difference was minimal (.23 points). The difference in mean PIQ was 11.79 points in the current study. Again, only one previous study of students with learning disabilities (Gridley, Arceneaux, Gayer, Buelow, & Palmer, 1994) had reported higher WISC-III than WISC-R PIQ means and the difference was considerably smaller (4.49 points). In addition, mean global scale scores for this sample of Native

Alaskan students with learning disabilities were somewhat lower than for other studies of students with learning disabilities.

The sample in the current study does differ from other samples on two key variables: ethnicity and community size. All students were Native Alaskan and living in small communities of less than 2500 in a rural isolated area of northern Alaska. While practice effects may account for some of the increases on the Performance scale subtests, it is unlikely they account for the entire change.

Mean verbal subtest scores on both the WISC-III and the WISC-R were in the below average area and characterized by a restricted range with standard deviations ranging from 1.73 to 2.48 on the WISC-III and from 1.47 to 2.43 on the WISC-R. More variability was noted on the Performance scale with mean subtest scores in the below average to average range with standard deviations ranging from 2.79 to 4.57 on the WISC-III and from 2.03 to 2.98 on the WISC-R. The difference in performance of these students is largely the result of the greater variability on the Performance subtests of the WISC-III. While mean differences on the Performance subtests were consistent (ranging from 1.17 to 2.10), variability in scores as shown by standard deviations in excess of 3.0 on the WISC-III was noted on Picture Completion, Picture Arrangement, Block Design, and Object Assembly.

Limitations of the current study include the relatively small sample size of 47. In addition, the sample represented only one ethnic group (Native Alaskan students) located in a rural, isolated area of northern Alaska. This study, as well as others, depended on previously administered WISC-Rs. Thus, practice effects cannot be ruled out. The alternative approach, administering both the WISC-III and WISC-R in counterbalanced order, would have reduced the

sample size and was not feasible. Additional research, utilizing this methodology is strongly encouraged, however.

Summary and Conclusions

The results of the current study suggest stability of performance in the verbal areas of the Wechsler with nonsignificant differences between WISC-III and WISC-R mean subtest scores on Information, Similarities, Arithmetic, Vocabulary, and Digit Span. VIQ scores were 4.64 points higher on the WISC-III. Scores on the Performance Scale of the WISC-III were characterized by a more variable level of performance with significant differences on the five Performance subtests (as compared to the previous WISC-R scores) and greater variability in subtest scores with standard deviations greater than 3.0 on four of the five Performance subtests. Whether these differences are related to the ethnicity variable or are artifacts of the current study is not clear. Further research comparing WISC-III/WISC-R performance for minority students with learning disabilities is needed to explore these issues.

The results of the current study in combination with previously cited studies indicate that:

1. The WISC-III produces VIQ and FSIQ scores that are consistent with those generated by the WISC-R. For this sample of Native Alaskan students with learning disabilities the WISC-III scores were higher than the WISC-R scores and PIQ scores were approximately 12 points higher on the WISC-III.
2. The correlational data for the WISC-III/WISC-R global scales suggest that they are measuring similar constructs (correlations ranged from .66 to .79). Subtest correlations were in the moderate range (.16 to .57).
3. Differences in scores are most likely to occur on the Performance Scale and least likely to occur on the Verbal Scale.

4. Largest mean differences on subtests (more than one point) occurred on Picture Completion, Picture Arrangement, Block Design, Object Assembly, and Coding.

5. Smallest mean differences on subtests (less than one-half point) occurred on Information, Arithmetic, and Vocabulary.

6. Use of the WISC-III rather than the WISC-R is strongly supported.

References

Bolen, L. M., Aichinger, K. S., Hall, C. W., & Webster, R. E. (1993, April). Comparison of the performance of learning impaired children on the WISC-R and WISC-III. Paper presented at Annual Meeting of National Association of School Psychologists, Washington, DC.

Graf, M. H., & Hinton, R. N. (1994). A three year comparison study of WISC-R and WISC-III IQ scores for a sample of special education students. Educational and Psychological Measurement, 54(1), 128-133.

Gridley, B. E., Arceneaux, J., Gayer, H., Buelow, B., & Palmer, C. (1994, August). Comparisons between WISC-R and WISC-III in reevaluation of special education students. Poster presented at the Annual Meeting of the American Psychological Association, Los Angeles.

Potter, A., & Wessel, J. (1993, October). Analysis of WISC-III data from an urban population of referred children. Paper presented at Fall Meeting of Wisconsin School Psychologists, Oshkosh, WI.

Slate, J. R., & Saarnio, D. A. (1995). Differences between WISC-III and WISC-R IQs; A preliminary investigation. Journal of Psychoeducational Assessment, 13(4), 340-346.

Smith, D. K., Stovall, D. L., & Geraghty, B. L. (1995). WISC-III/WISC-R relationships in special education re-evaluations. The Wisconsin School Psychologist, 24(3), 1-6.

Weiss, L. G. (1995). WISC-III IQs: New norms raise queries: A review of 22 studies comparing WISC-R to WISC-III scores. Assessment Focus, 1(2), 1, 4-6.

Wechsler, D. W. (1991). Wechsler Intelligence Scale for Children-Third Edition. San Antonio, TX: The Psychological Corporation.

Table 1

Comparison of WISC-III and WISC-R IQ Scores

Scale	n		WISC-III	WISC-R	Mean Difference	t	p
Verbal Scale	47	M	75.66	71.02	4.64	4.90	<.001
		SD	(10.43)	(9.02)			
Performance Scale	47	M	92.47	80.68	11.79	6.44	<.001
		SD	(16.56)	(12.71)			
Full Scale	47	M	81.72	73.72	8.00	6.78	<.001
		SD	(12.88)	(9.87)			

Table 2

Comparison of WISC-III and WISC-R Subtest Scores

Subtests	n		WISC-III	WISC-R	Mean	t	p
Information	47	M	5.70	5.21	+.49	+1.70	NS
		SD	1.78	1.47			
Similarities	47	M	6.26	5.38	+.88	+2.00	NS
		SD	2.06	2.43			
Arithmetic	47	M	5.10	5.57	-.47	-1.61	NS
		SD	1.97	2.23			
Vocabulary	47	M	5.19	5.38	-.19	-2.00	NS
		SD	1.73	2.43			
Comprehension	47	M	6.04	5.14	+.90	+2.43	<.05
		SD	2.48	2.01			
(Digit Span)	39	M	6.00	5.19	+.81	+1.40	NS
		SD	2.42	2.07			
Picture Completion	47	M	9.06	7.89	+1.17	+2.62	<.05
		SD	3.82	2.03			
Picture Arrangement	47	M	7.27	5.91	+1.36	+2.65	<.05
		SD	3.38	2.67			
Block Design	47	M	9.00	6.90	+2.10	+3.02	<.01
		SD	4.57	2.98			
Object Assembly	47	M	9.21	7.72	+1.49	+3.38	<.001
		SD	3.74	2.34			
Coding	46	M	8.53	7.11	+1.42	+2.63	<.05
		SD	2.79	2.62			

Table 3

WISC-III and WISC-R Correlations

Global Scales	**n**	**Pearson Correlation**	**p**
Verbal Scale	47	.79	<.001
Performance	47	.66	<.001
Full Scale	47	.76	<.001
Subtests			
Information	47	.26	NS
Similarities	47	.20	NS
Arithmetic	47	.57	<.001
Vocabulary	47	.32	<.03
Comprehension	47	.52	<.001
(Digit Span)	46	.36	<.05
Picture Completion	47	.36	<.05
Picture Arrangement	47	.16	NS
Block Design	47	.31	<.05
Object Assembly	47	.44	<.01
Coding	47	.41	<.01